NASA Data Products for HELIX-Atlanta

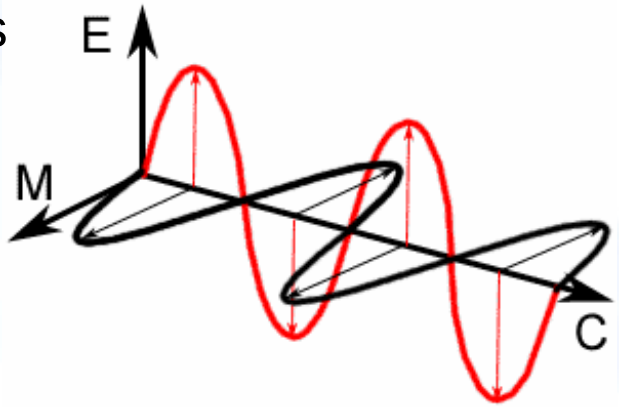
Doug Rickman, Ph.D. and Dale Quattrochi, Ph.D.
NASA
Marshall Space Flight Center
Huntsville, Alabama

Learning Objectives

- How NASA data products for HELIX-Atlanta are acquired, stored and shipped
- Standards for data products for HELIX-Atlanta
- Value added to HELIX-Atlanta by including NASA data products

The Source

- NASA has 18 satellites which continuously observe the third planet in the solar system. Each satellite carries one or more sensors, measuring energies across the electromagnetic spectrum.



- Data from the sensors is beamed to the Earth, archived, manipulated, and distributed on a continuous basis. From the raw data specialized products are made, quantifying and mapping variables of interest at spatial scales ranging from 100+ Km to meters.

The Scale

- These data are in one of the largest digital archives in the world. Last year over 200,000+ users obtained pieces of this unparalleled record of the Earth's geophysical and environmental history, state, and dynamics.
- To make this happen NASA and partner agencies maintain large staffs dedicated to providing the more than 10,000,000 products shipped each year.
- Utilizing a suite of Distributed Active Archive Centers (DAAC), data sets and products are stored topically. Users may obtain data a number of ways, including web down load, ftp, and tape. Data are generally in what is called the Earth Observing System (EOS), Hierarchical Data Format (HDF) or EOS-HDF format, a public specification useful for large data sets with significant quantities of metadata.
- These data are all a part of what is called the Earth Observing System Distributed Information System or EOSDIS

Who

The following information is from "EOSDIS Users and Usage - What We Know About Our Users"; Vanessa Griffin, Jeanne Behnke, Robert Wolfe, Kathy Fontaine, Steve Adamson; Presentation to ESISS, Scripps Institute, February 17, 2004

EOSDIS Users in FY2003

- The users that received data from other ESE-funded data systems, including flight project web sites, SIPS, SCFs, ESIPs, REASoNs, and Direct Broadcast sites, are not included in this study.
- 228,000 distinct users received data and information products from EOSDIS in FY2003
  - 17,000 users placed orders for NASA's Earth science data products using the various user interfaces available at the eight DAACs
  - Remainder of users downloaded data products via FTP or information products via the web
  - This does not include users who ordered ECS data products from the new EOS data pools
  - 78% of these users were "first time" users (an email address that was seen for the first time in 2003)

EOSDIS Usage in FY2003

- EOSDIS users received more than 29 million data and information products.
  - 15 million were data products delivered per a specific user request
  - 11 million were “on-line” data and information products retrieved from anonymous FTP servers at DAACs
  - 3 million information products were download from DAAC web sites
- There was a consistent growth in the average number of products requested per order over the past four years, possibly due to:
  - increased use of data for time series applications, or
  - growing familiarity with and/or improvements to product ordering interfaces.

What the ECS Data Usage Metrics Show

- 40% of users were from countries other than the USA
- The number of ECS products requested has steadily increased over the past four years
- Increasingly, users are ordering higher level products (L2-L4)
  - 70% of users received a higher level product in FY2003
  - Four year trends show that users in all domains are requesting more higher level data products (Levels 2-4)
  - US Government users are the highest percentage of users of lower level data, but they, too, are ordering increasing amounts of level 2 and 3 data products
  - Usage by university researchers (.edu) is predominately of Level 2 and Level 3 data products

Top EOSDIS Products

(by number of times product was distributed)

ECS Data Products

1. MODIS Radiances (L1B)
2. Landsat 7 Radiances (L0R)
3. ASTER Surface Reflect (L2)
4. MODIS Cloud Mask (L2)
5. ASTER DEM (Level 4)
6. MODIS Aerosols (L2)
7. MODIS Cloud Product
8. ASTER Radiances (L1B)
9. MODIS Snow Cover (L3)
10. MODIS Atmospheres (L3)
11. ASTER Surf Emissivity (L2)
12. ASTER Surf Temp (L2)
13. MODIS Ocean Color (L2)

Other Earth Science Products

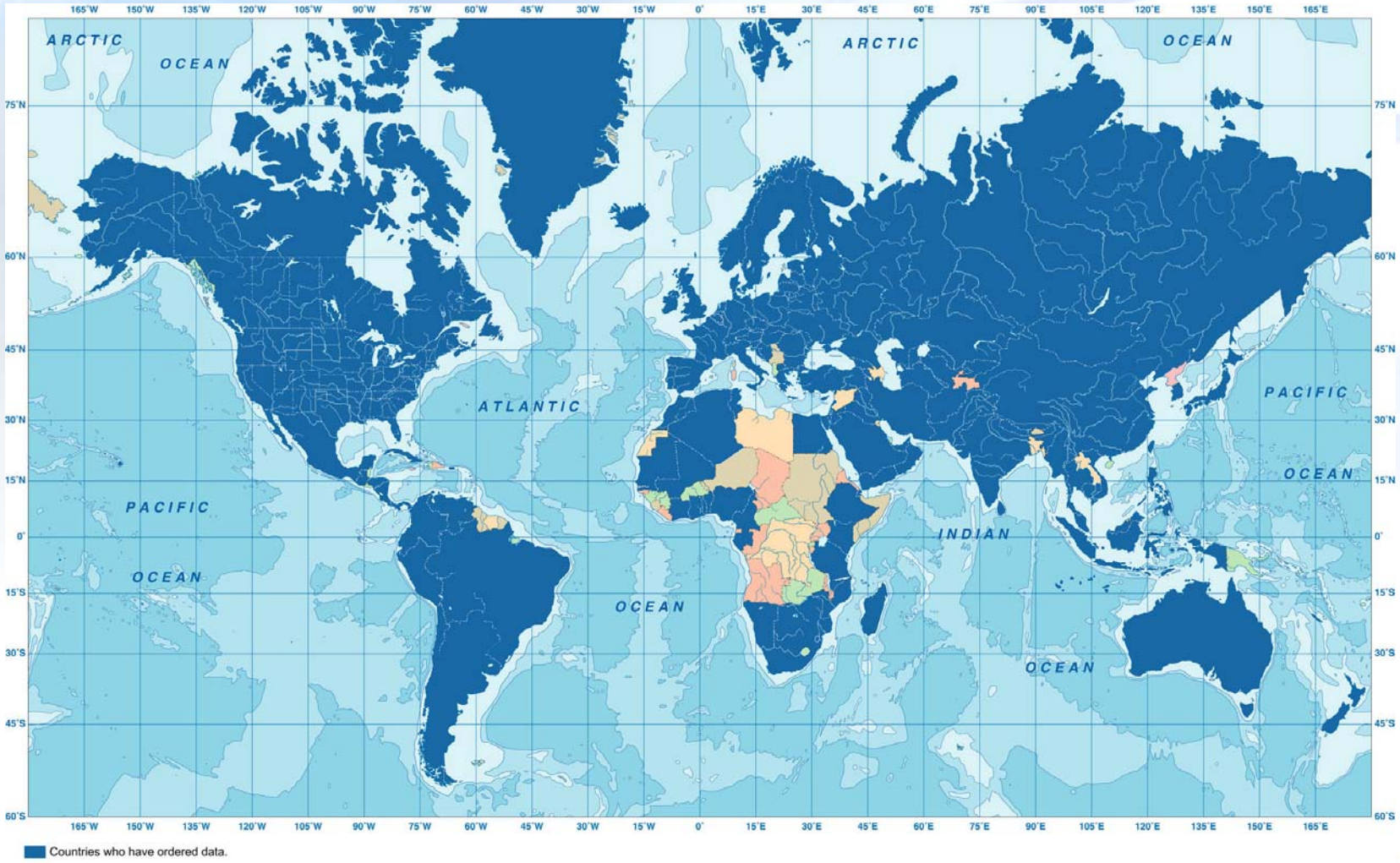
1. TOPEX/P Sfc Height anom (L3)
2. SSMI Brightness Temp (L2)
3. NWS RADAR Reflec
4. TOMS Total Ozone (L3)
5. NCEP Model Output (L4)
6. SEAWiFS Image Data (L1)
7. QuikSCAT Vectors (L2)
8. TRMM Bright Temp (l2)
9. TRMM PR Rain Profiles (L2)
10. JASON Raw Data Records
11. AVHRR SST (l3)

- These products only represent those distributed by the DAACs and do not include the distribution by SIPS, SCFs, Projects, ESIPs, etc.
- While 70% of ECS users are ordering Level 2 and higher products, large numbers of lower level radiance products continue to be requested

What the Broader Usage Statistics Show

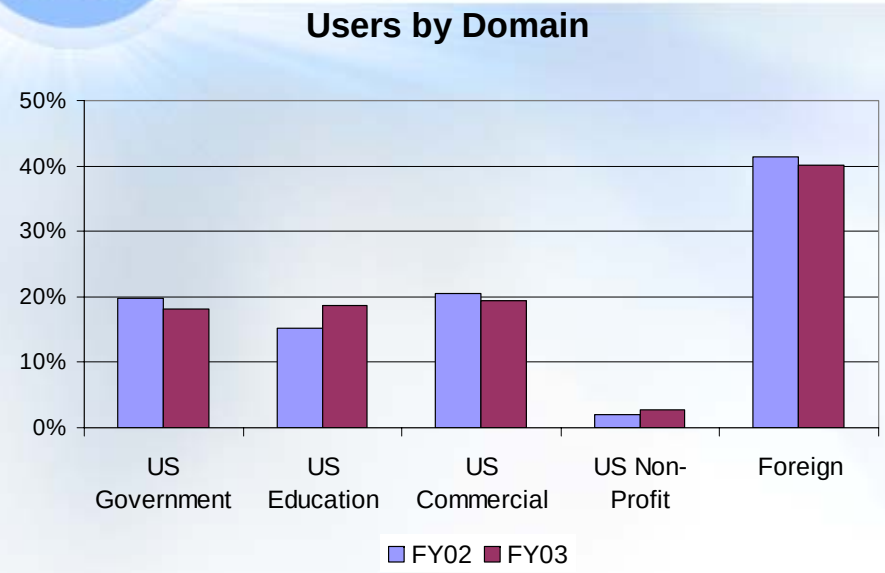
- More and more, users are downloading data electronically via networks.
  - There has been a dramatic increase over four years in the number or products retrieved immediately from web sites, FTP servers and the Data Pools
  - 58% of all products were delivered immediately in FY2003 compared to just 27% in FY2000
  - “Staged” electronic deliveries increased from 1.5% to 19%
  - Media (tapes, CD-ROMS) deliveries fell from 72% to 23% over the period
- Overall distribution of Earth science products to end users (not counting SCF and SIPS) grew at 10-20% per year between FY2000 and FY2002

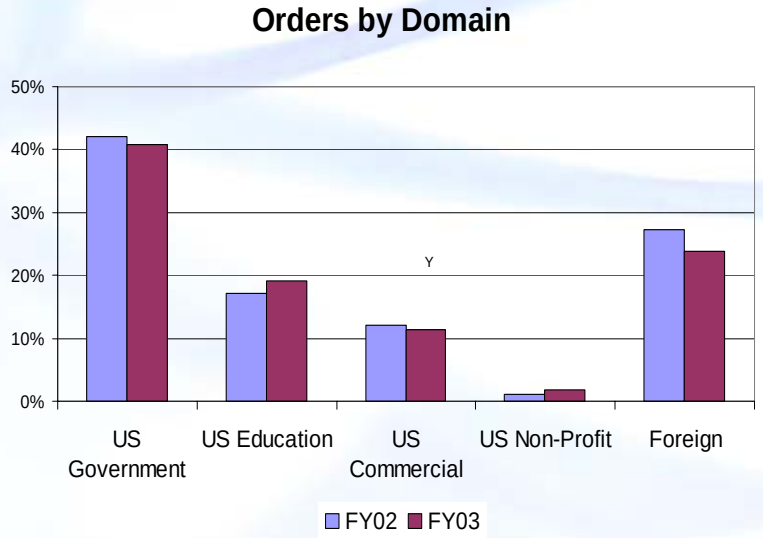
Countries Ordering Data

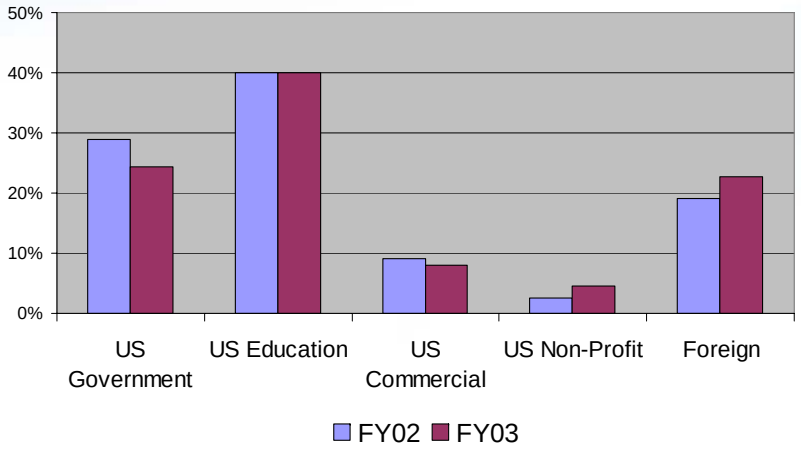


In 2003 products were distributed to 139 countries

ECS User Trends (by domain)

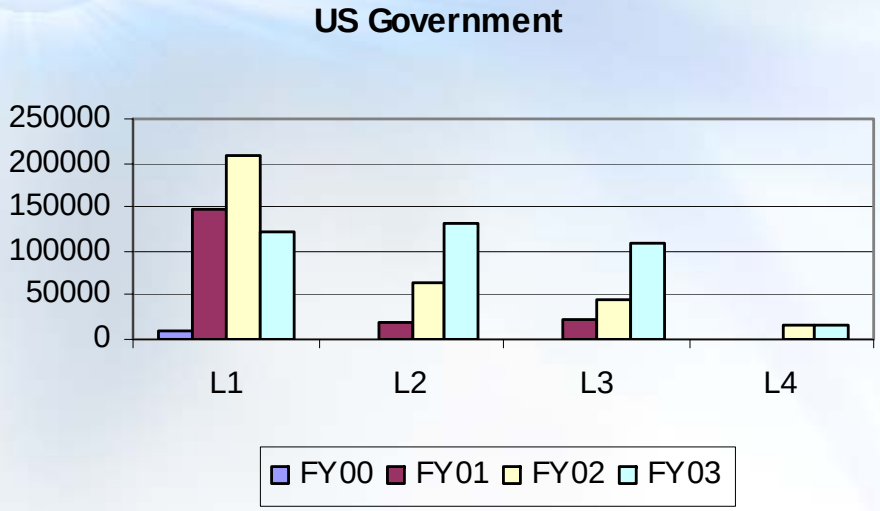


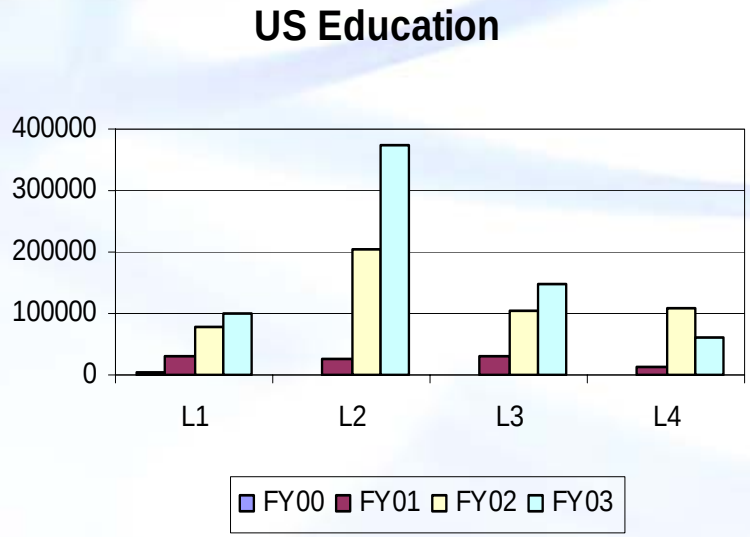


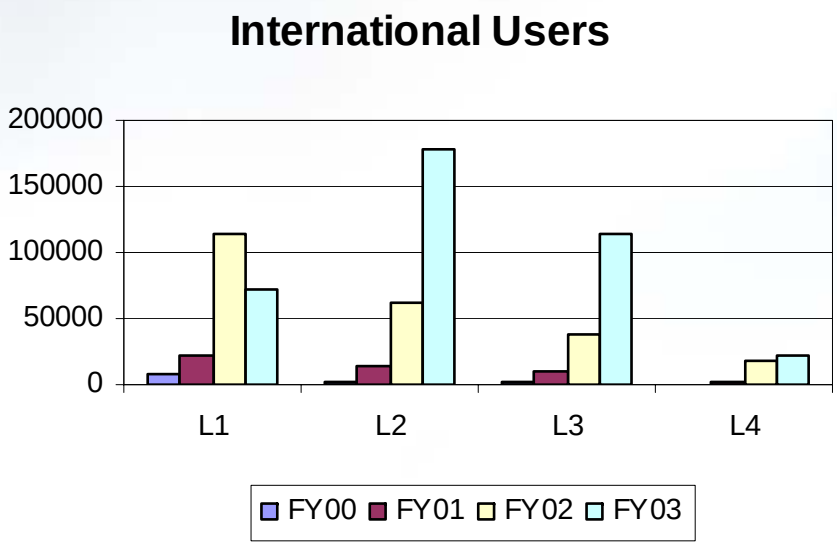


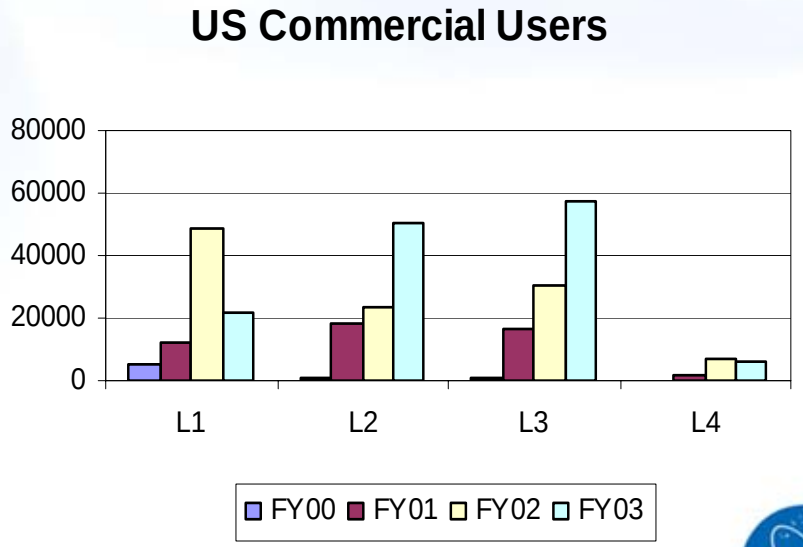
- 40% of users were international
- Over 40% of all products were requested by university users

ECS Products Delivered (by level) Per Domain

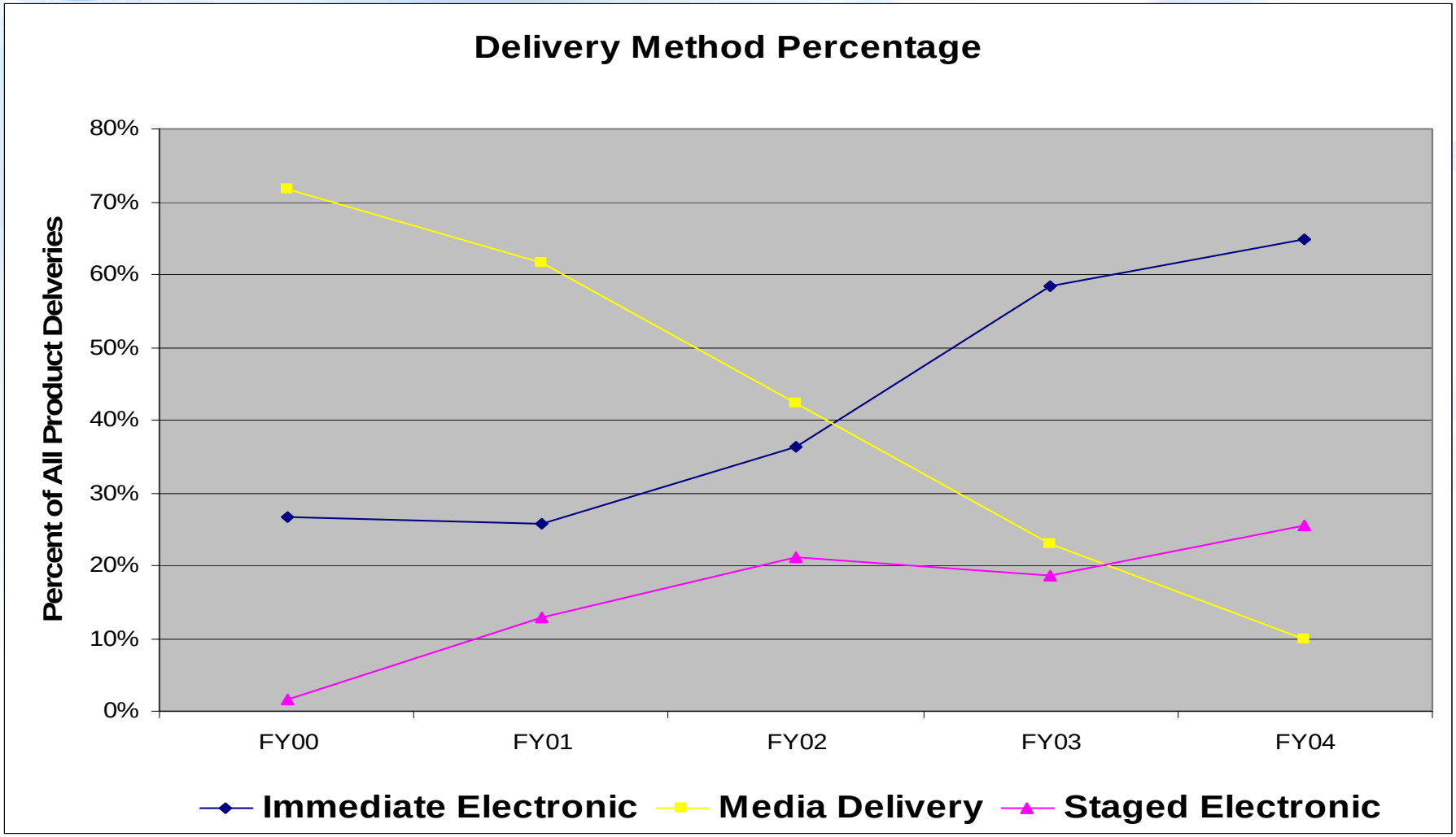








Trends in Delivery Options By User Domain



DAAC Evolution

- More and more users need access to near real time data.
  - *ESE needs to focus on increasing timeliness of data products*
- Users have problems downloading large volumes of data (especially when they only need a few “bytes” from a file)
  - This is critical for researchers in the “field”
  - DAACs have implemented subsetting approaches to help users
  - *ESE should look at providing more services designed to help users locate and retrieve only the minimal amount data they need for their application*
- EOS geophysical parameter algorithms are not always optimal for specific regional areas or research applications
  - Researchers order Level 1 radiance data and apply their own optimized algorithms
  - *ESE will need to continue serving lower level data products to users*

DAAC Evolution

- EOS standard data products are primarily oriented toward the Earth science researcher community and need some customization before they can become useful to other user groups
- The on-line data pools were praised as a good data distribution method.
- NASA data (i.e., EOS-HDF formatted data) will be in compliance with the emerging National Spatial Data Infrastructure (NSDI) data standards that are being developed for all spatial data distributed by U.S. federal agencies.
- The difficulty of co-locating and inter-comparing datasets from different instruments and sometimes across disciplines is a barrier to the use of the data.
- Data persistence is a key concern particularly for the education community.

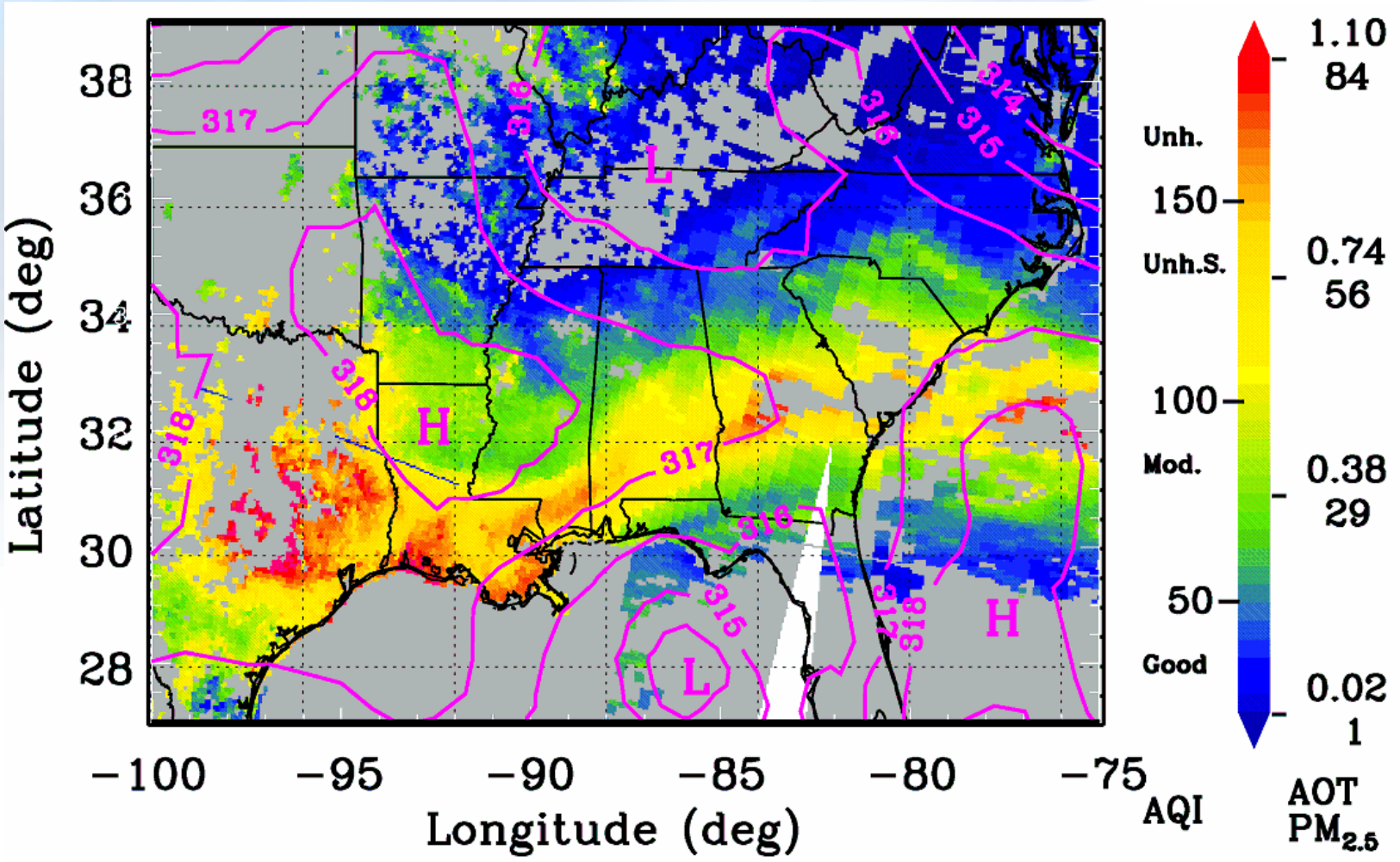
Barriers to Fullest Use by Operational Community

- Consistency in data quality (i.e., cross-calibration, validation of products).
- Timeliness of data delivery (this is going to be application dependent, but one operational user recommended delivery of land products in the same time frame as weather products).
- Data formats that are compatible with the operational community needs (e.g., GIS and NSDI standards)
- Provide services and tools to make EOS-HDF data available in alternative formats such as GeoTIFF, NetCDF, and JPEG

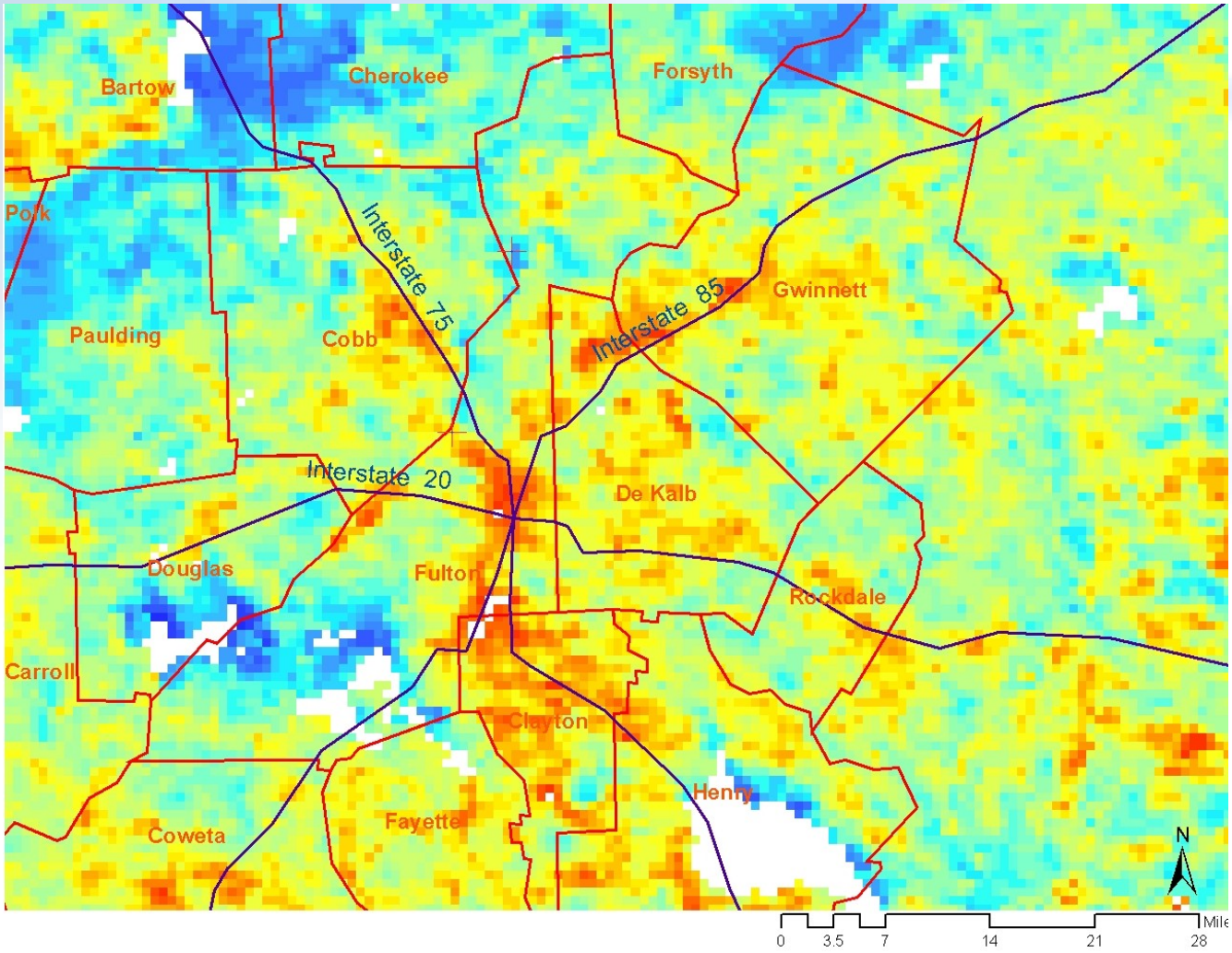
NASA data for HELIX-Atlanta

- From the available data we have selected several candidates for use within HELIX-Atlanta: Surface Temperature, PM2.5, and land usage.
- The following are for illustration. Specific implementation will likely have variation from these.

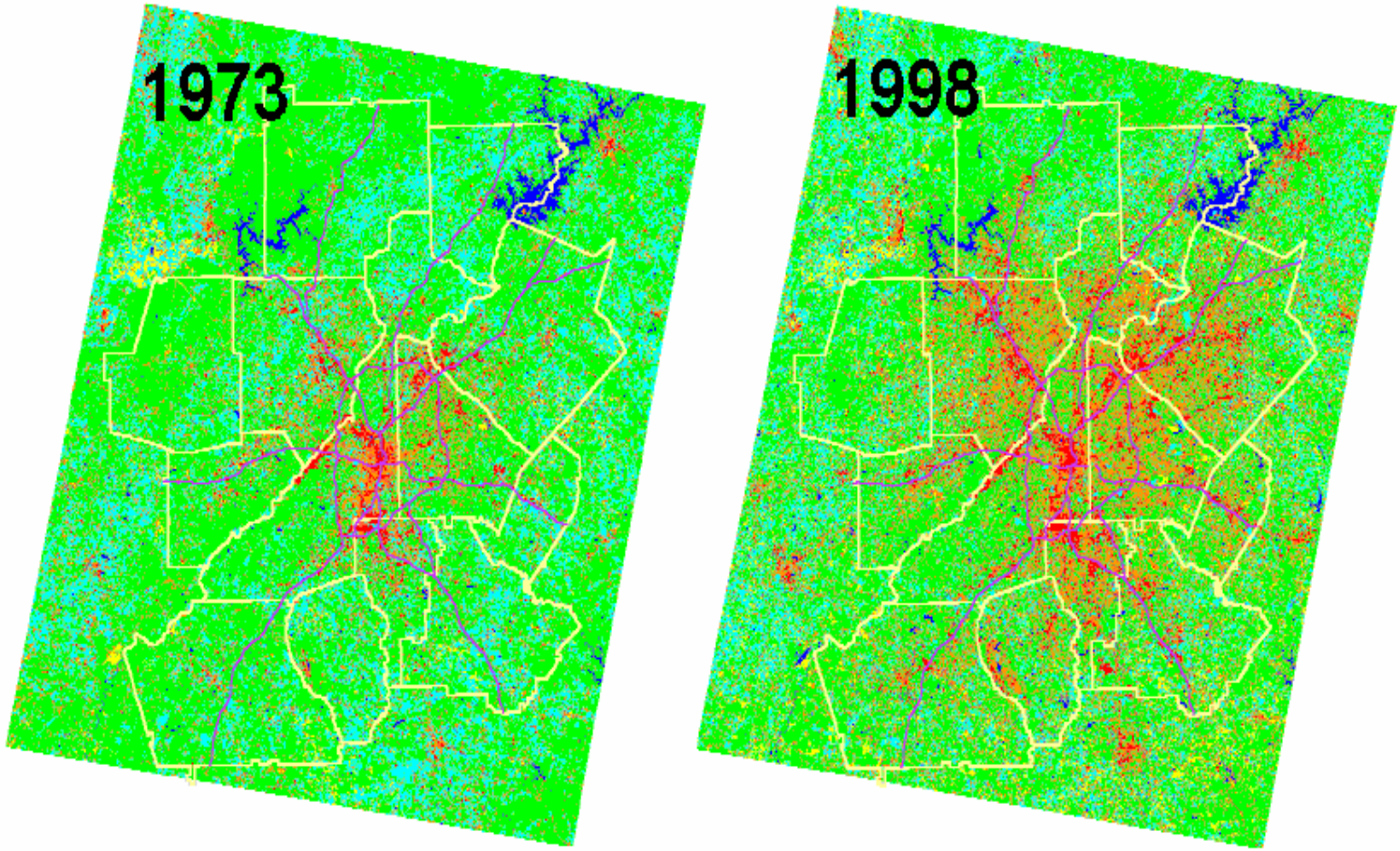
PM2.5 (inferred from MODIS data)



Surface Temperature (MODIS)



Land Use Classification (hypothetical)



Considerations for PHIN

- Bringing these types of data into PHIN presents many problems in system design. Some of the points we have already recognized are frequency and mode of access, payment if necessary, copy rights and permissions, data volume for handling and storage, and GIS standards.
- These data, however, provide an unmatched or unparallel view of environmental variables.

Our Contact Information

- Web Site: http://www.ghcc.msfc.nasa.gov/
- E-mail addresses and telephone numbers:
  - Doug Rickman: doug.rickman@nasa.gov
    - Phone: (256) 961-7889
  - Dale Quattrochi: dale.quattrochi@nasa.gov
    - Phone: (256) 961-7887